

OIPE

RAW SEQUENCE LISTING

DATE: 09/21/2001

PATENT APPLICATION: US/09/945,254

TIME: 20:48:55

Input Set : A:\Seqlist.txt

Output Set: N:\CRF3\09212001\I945254.raw

ENTERED

```

4 <110> APPLICANT: Meyers, Rachel
5      MacBeth, Kyle
6      Tsai, Fong-Ying
8 <120> TITLE OF INVENTION: 8797, A NOVEL HUMAN
9      GALACTOSYLTRANSFERASE AND USES THEREOF
11 <130> FILE REFERENCE: MNI-188
C--> 13 <140> CURRENT APPLICATION NUMBER: US/09/945,254
C--> 13 <141> CURRENT FILING DATE: 2001-08-31
13 <150> PRIOR APPLICATION NUMBER: 60/229,829
14 <151> PRIOR FILING DATE: 2000-08-31
16 <160> NUMBER OF SEQ ID NOS: 3
18 <170> SOFTWARE: FastSEQ for Windows Version 4.0
20 <210> SEQ ID NO: 1
21 <211> LENGTH: 4052
22 <212> TYPE: DNA
23 <213> ORGANISM: Homo sapiens
25 <220> FEATURE:
26 <221> NAME/KEY: CDS
27 <222> LOCATION: (459)...(1592)
29 <400> SEQUENCE: 1
30 ccaagattta aagcccgc aa gttttgttct tgagaccagc gacttttagct ccgatgcggg 60
31 aaggaaagcc gacctccgat ttggacattt aaagagctgg gcttgaactt cgtgagtttc 120
32 gctctaaact gcccttgaaa tgaagctgga cttggagggtg gcatggaata ttcacatggg 180
33 agagccgcat gaggccgccc accacgcttc ctgaaggatg cccgtgtgga agaattttga 240
34 cgtgccagtg tctcgttct acagggtgtt ccattcttcc gcaatctcag aaaaatggga 300
35 ctaaaagaaa ctattttgta aaataagaag acttccattt ttaatgacca acatgtatta 360
36 agatggacac ctactctacg aaacacgaag ttctatggtc tcgaagaagc ccgtgcctgt 420
37 ttaaaactga tcctaactaa aaacagactt gagtggat atg aga atg ttg gtt agt 476
38                                     Met Arg Met Leu Val Ser
39                                     1           5
41 ggc aga aga gtc aaa aaa tgg cag tta att att cag tta ttt gct act 524
42 Gly Arg Arg Val Lys Lys Trp Gln Leu Ile Ile Gln Leu Phe Ala Thr
43          10           15           20
45 tgt ttt tta gcg agc ctc atg ttt ttt tgg gaa cca atc gat aat cac 572
46 Cys Phe Leu Ala Ser Leu Met Phe Phe Trp Glu Pro Ile Asp Asn His
47          25           30           35
49 att gtg agc cat atg aag tca tat tct tac aga tac ctc ata aat agc 620
50 Ile Val Ser His Met Lys Ser Tyr Ser Tyr Arg Tyr Leu Ile Asn Ser
51          40           45           50
53 tat gac ttt gtg aat gat acc ctg tct ctt aag cac acc tca gcg ggg 668
54 Tyr Asp Phe Val Asn Asp Thr Leu Ser Leu Lys His Thr Ser Ala Gly
55          55           60           65           70
57 cct cgc tac caa tac ttg att aac cac aag gaa aag tgt caa gct caa 716
58 Pro Arg Tyr Gln Tyr Leu Ile Asn His Lys Glu Lys Cys Gln Ala Gln
59          75           80           85
61 gac gtc ctc ctt tta ctg ttt gta aaa act gct cct gaa aac tat gat 764
62 Asp Val Leu Leu Leu Leu Phe Val Lys Thr Ala Pro Glu Asn Tyr Asp

```

RAW SEQUENCE LISTING

DATE: 09/21/2001

PATENT APPLICATION: US/09/945,254

TIME: 20:48:55

Input Set : A:\Seqlist.txt

Output Set: N:\CRF3\09212001\I945254.raw

63	90	95	100	
65	cga cgt tcc gga att aga agg acg tgg ggc aat gaa aat tat gtt cgg	812		
66	Arg Arg Ser Gly Ile Arg Arg Thr Trp Gly Asn Glu Asn Tyr Val Arg			
67	105	110	115	
69	tct cag ctg aat gcc aac atc aaa act ctg ttt gcc tta gga act cct	860		
70	Ser Gln Leu Asn Ala Asn Ile Lys Thr Leu Phe Ala Leu Gly Thr Pro			
71	120	125	130	
73	aat cca ctg gag gga gaa gaa cta caa aga aaa ctg gct tgg gaa gat	908		
74	Asn Pro Leu Glu Gly Glu Glu Leu Gln Arg Lys Leu Ala Trp Glu Asp			
75	135	140	145	150
77	caa agg tac aat gat ata att cag caa gac ttt gtt gat tct ttc tac	956		
78	Gln Arg Tyr Asn Asp Ile Ile Gln Gln Asp Phe Val Asp Ser Phe Tyr			
79	155	160	165	
81	aat ctt act ctg aaa tta ctt atg cag ttc agt tgg gca aat acc tat	1004		
82	Asn Leu Thr Leu Lys Leu Leu Met Gln Phe Ser Trp Ala Asn Thr Tyr			
83	170	175	180	
85	tgt cca cat gcc aaa ttt ctt atg act gct gat gat gac ata ttt att	1052		
86	Cys Pro His Ala Lys Phe Leu Met Thr Ala Asp Asp Asp Ile Phe Ile			
87	185	190	195	
89	cac atg cca aat ctg att gag tac ctt caa agt tta gaa caa att ggt	1100		
90	His Met Pro Asn Leu Ile Glu Tyr Leu Gln Ser Leu Glu Gln Ile Gly			
91	200	205	210	
93	gtt caa gac ttt tgg att ggt cgt gtt cat cgt ggt gcc cct ccc att	1148		
94	Val Gln Asp Phe Trp Ile Gly Arg Val His Arg Gly Ala Pro Pro Ile			
95	215	220	225	230
97	aga gat aaa agc agc aaa tac tac gtg tcc tat gaa atg tac cag tgg	1196		
98	Arg Asp Lys Ser Ser Lys Tyr Tyr Val Ser Tyr Glu Met Tyr Gln Trp			
99	235	240	245	
101	cca gct tac cct gac tac aca gcc gga gct gcc tat gta atc tcc ggt	1244		
102	Pro Ala Tyr Pro Asp Tyr Thr Ala Gly Ala Ala Tyr Val Ile Ser Gly			
103	250	255	260	
105	gat gta gct gcc aaa gtc tat gag gca tca cag aca cta aat tca agt	1292		
106	Asp Val Ala Ala Lys Val Tyr Glu Ala Ser Gln Thr Leu Asn Ser Ser			
107	265	270	275	
109	ctt tac ata gac gat gtg ttc atg ggc ctc tgt gcc aat aaa ata ggg	1340		
110	Leu Tyr Ile Asp Asp Val Phe Met Gly Leu Cys Ala Asn Lys Ile Gly			
111	280	285	290	
113	ata gta ccg cag gac cat gtg ttt ttt tct gga gag ggt aaa act cct	1388		
114	Ile Val Pro Gln Asp His Val Phe Phe Ser Gly Glu Gly Lys Thr Pro			
115	295	300	305	310
117	tat cat ccc tgc atc tat gaa aaa atg atg aca tct cat gga cac tta	1436		
118	Tyr His Pro Cys Ile Tyr Glu Lys Met Met Thr Ser His Gly His Leu			
119	315	320	325	
121	gaa gat ctc cag gac ctt tgg aag aat gct aca gat cct aaa gta aaa	1484		
122	Glu Asp Leu Gln Asp Leu Trp Lys Asn Ala Thr Asp Pro Lys Val Lys			
123	330	335	340	
125	acc att tcc aaa ggt ttt ttt ggt caa ata tac tgc aga tta atg aag	1532		
126	Thr Ile Ser Lys Gly Phe Phe Gly Gln Ile Tyr Cys Arg Leu Met Lys			
127	345	350	355	

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/945,254

DATE: 09/21/2001

TIME: 20:48:55

Input Set : A:\Seqlist.txt

Output Set: N:\CRF3\09212001\I945254.raw

```

129 ata att ctc ctt tgt aaa att agc tat gtg gac aca tac cct tgt agg 1580
130 Ile Ile Leu Leu Cys Lys Ile Ser Tyr Val Asp Thr Tyr Pro Cys Arg
131      360                      365                      370
133 gct gcg ttt atc taatagtact tgaatgttgt atgttttcac tgtcactgag 1632
134 Ala Ala Phe Ile
135 375
137 tcaaacctgg atgaaaaaaaa cctttaaatg ttogtctata ccctaagtaa aatgaggacg 1692
138 aaagacaaat attttgaaag cctagtccat cagaatgttt ctttgattct agaagctgtt 1752
139 taatatcact tatctacttc attgcctaag ttcatttcaa agaatttgta tttagaaaaa 1812
140 gtttatatta ttagtgaaaa caaaactaaa gggaagtcca agttctcatg taatgccaca 1872
141 tatatacttg aggtgtagag atgttattaa gaagttttga tgttagaata attgcttttg 1932
142 gaaaataacca aatgaacgta cagtacaaca tttcaaggaa atgaatatat tgtagacca 1992
143 ggtaagcaag tttatttttg ttaaagagca cttgggtggag gtagtagggg cagggaaaag 2052
144 tcagcatagg agagaaagt catgaatctg gtaaaacagt ctcttggtct taagaggaga 2112
145 tgtagaaaaa tgtgtacaat gttattataa acagacaaat cagctcttac cacatccatg 2172
146 tagctactgg tgtagagtc attaaaaata ctttttttgc atcttttttc aaagttaa 2232
147 gtgaactttt agaaaagtga ttaattgtgc cctaatactt tatatgtttt taatggattt 2292
148 ttttttaagt attagaaaat gacacataac acgggcagct ggttgctcat agggctcttc 2352
149 tctagggaga aaccattggt aattcaaata agctgatttt aatgacgttt tcaactggtt 2412
150 tttaaatatt caatattggt ctgtgtttta gtttgttatt tgaatgtaat ttacatagag 2472
151 gaatataata atggagagac ttcaaattgga aagacagaac attacaagcc taatgtctcc 2532
152 ataattttat aaaatgaaat cttagtgtct aaatccttgt actgattact aaaattaacc 2592
153 cactcctccc caacaaggtc ttataaacca cagcactttg ttccaagttc agagttttaa 2652
154 attgagagca ttaaacatca aagtataat atctaaaaca atttattttt catcaataac 2712
155 tgtcagaggt gatctttatt ttctaaatat ttcaaacttg aaaacagagt aaaaaagtga 2772
156 tagaaaagtt gccagtttgg ggtaaagca tttttaaagc tgcattgtcc ttgtaataca 2832
157 agagatgtgt ctgagatcta atagagtaag ttacatttat tttaaaaagc aggataaaaa 2892
158 tgtggctata atacacacta cctcccttca ctacagaaag aactagggtg tgtctactgc 2952
159 tagggagatt atatgaaggc caaaataatg acttcagcaa gactgactga actcactcta 3012
160 aggcctttga ctgcagaggc acctgttagg gaaaatcaga tgtctcatat aataaggtga 3072
161 tgtcggaaac atcgcaaaac aaacgaaaaa agatttctca gtatacaca ctgaatgatg 3132
162 atacttacaa tttttagcag gtagcttttt aatgtttaca gaaattttta tttttttcta 3192
163 ttttgaaatt tgaggcttgt ttacattgct tagataattt agaattttta actaatgtca 3252
164 aaactacagt gtcaaacatt ctagggttga gttactttca gagtagatac agggtttttag 3312
165 atcattacag ttttaagtgt ctgaccaatt aaaaaaacat agagaacaaa agcatatttg 3372
166 accaagcaac aagcttataa ttaattttta ttagttgatt gattaatgat gtattgcctt 3432
167 ttgcccatat atacctgtg tatctatact tggaagtgtt taagggtgcc attggttgaa 3492
168 aacataagtg tctctggcca tcaaagtgat cttgtttaca gcagtgcctt tgtgaaacaa 3552
169 ttattttatt gctgaaaagag ctcttctgaa ctgtgtcctt ttaatttttg cttagaatag 3612
170 aatggaacaa gtttaaattt caaggaaata tgaaggcact tccttttttt ctaagaagga 3672
171 agttgctaga tgattccttc atcacactta cttaaagtac tgagaagagt atctgtaaat 3732
172 aaaagggttc caacctttta aaaaagaagg aaaaaacttt ttggtgctcc agttagaggc 3792
173 tatcttttta aaaaatgtca acaaagggaa aataaactat cagcttggat ggtcacttga 3852
174 atagaagatg gttatacaca gtgttattgt taaaattttt ttaccttttg gttggtttgc 3912
175 atcttttttc catattgtta attttatacc aaaatgttaa atatttgtat tacttgaatt 3972
176 ttgctcttgt atggcaaaat aattagttag tttaaaaaaa atctatagtt tccaataaac 4032
177 aactgaaaaa ttaaaaaaaa 4052
179 <210> SEQ ID NO: 2
180 <211> LENGTH: 378

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/945,254

DATE: 09/21/2001

TIME: 20:48:55

Input Set : A:\Seqlist.txt

Output Set: N:\CRF3\09212001\I945254.raw

```

181 <212> TYPE: PRT
182 <213> ORGANISM: Homo sapiens
184 <400> SEQUENCE: 2
185 Met Arg Met Leu Val Ser Gly Arg Arg Val Lys Lys Trp Gln Leu Ile
186 1 5 10 15
187 Ile Gln Leu Phe Ala Thr Cys Phe Leu Ala Ser Leu Met Phe Phe Trp
188 20 25 30
189 Glu Pro Ile Asp Asn His Ile Val Ser His Met Lys Ser Tyr Ser Tyr
190 35 40 45
191 Arg Tyr Leu Ile Asn Ser Tyr Asp Phe Val Asn Asp Thr Leu Ser Leu
192 50 55 60
193 Lys His Thr Ser Ala Gly Pro Arg Tyr Gln Tyr Leu Ile Asn His Lys
194 65 70 75 80
195 Glu Lys Cys Gln Ala Gln Asp Val Leu Leu Leu Phe Val Lys Thr
196 85 90 95
197 Ala Pro Glu Asn Tyr Asp Arg Arg Ser Gly Ile Arg Arg Thr Trp Gly
198 100 105 110
199 Asn Glu Asn Tyr Val Arg Ser Gln Leu Asn Ala Asn Ile Lys Thr Leu
200 115 120 125
201 Phe Ala Leu Gly Thr Pro Asn Pro Leu Glu Gly Glu Glu Leu Gln Arg
202 130 135 140
203 Lys Leu Ala Trp Glu Asp Gln Arg Tyr Asn Asp Ile Ile Gln Gln Asp
204 145 150 155 160
205 Phe Val Asp Ser Phe Tyr Asn Leu Thr Leu Lys Leu Leu Met Gln Phe
206 165 170 175
207 Ser Trp Ala Asn Thr Tyr Cys Pro His Ala Lys Phe Leu Met Thr Ala
208 180 185 190
209 Asp Asp Asp Ile Phe Ile His Met Pro Asn Leu Ile Glu Tyr Leu Gln
210 195 200 205
211 Ser Leu Glu Gln Ile Gly Val Gln Asp Phe Trp Ile Gly Arg Val His
212 210 215 220
213 Arg Gly Ala Pro Pro Ile Arg Asp Lys Ser Ser Lys Tyr Tyr Val Ser
214 225 230 235 240
215 Tyr Glu Met Tyr Gln Trp Pro Ala Tyr Pro Asp Tyr Thr Ala Gly Ala
216 245 250 255
217 Ala Tyr Val Ile Ser Gly Asp Val Ala Ala Lys Val Tyr Glu Ala Ser
218 260 265 270
219 Gln Thr Leu Asn Ser Ser Leu Tyr Ile Asp Asp Val Phe Met Gly Leu
220 275 280 285
221 Cys Ala Asn Lys Ile Gly Ile Val Pro Gln Asp His Val Phe Phe Ser
222 290 295 300
223 Gly Glu Gly Lys Thr Pro Tyr His Pro Cys Ile Tyr Glu Lys Met Met
224 305 310 315 320
225 Thr Ser His Gly His Leu Glu Asp Leu Gln Asp Leu Trp Lys Asn Ala
226 325 330 335
227 Thr Asp Pro Lys Val Lys Thr Ile Ser Lys Gly Phe Phe Gly Gln Ile
228 340 345 350
229 Tyr Cys Arg Leu Met Lys Ile Ile Leu Leu Cys Lys Ile Ser Tyr Val
230 355 360 365

```

RAW SEQUENCE LISTING

DATE: 09/21/2001

PATENT APPLICATION: US/09/945,254

TIME: 20:48:55

Input Set : A:\Seqlist.txt

Output Set: N:\CRF3\09212001\I945254.raw

```

231 Asp Thr Tyr Pro Cys Arg Ala Ala Phe Ile
232      370                      375
235 <210> SEQ ID NO: 3
236 <211> LENGTH: 1134
237 <212> TYPE: DNA
238 <213> ORGANISM: Homo sapiens
240 <220> FEATURE:
241 <221> NAME/KEY: CDS
242 <222> LOCATION: (1)...(1134)
244 <400> SEQUENCE: 3
245 atg aga atg ttg gtt agt ggc aga aga gtc aaa aaa tgg cag tta att 48
246 Met Arg Met Leu Val Ser Gly Arg Arg Val Lys Lys Trp Gln Leu Ile
247 1      5      10      15
249 att cag tta ttt gct act tgt ttt tta gcg agc ctc atg ttt ttt tgg 96
250 Ile Gln Leu Phe Ala Thr Cys Phe Leu Ala Ser Leu Met Phe Phe Trp
251      20      25      30
253 gaa cca atc gat aat cac att gtg agc cat atg aag tca tat tct tac 144
254 Glu Pro Ile Asp Asn His Ile Val Ser His Met Lys Ser Tyr Ser Tyr
255      35      40      45
257 aga tac ctc ata aat agc tat gac ttt gtg aat gat acc ctg tct ctt 192
258 Arg Tyr Leu Ile Asn Ser Tyr Asp Phe Val Asn Asp Thr Leu Ser Leu
259      50      55      60
261 aag cac acc tca gcg ggg cct cgc tac caa tac ttg att aac cac aag 240
262 Lys His Thr Ser Ala Gly Pro Arg Tyr Gln Tyr Leu Ile Asn His Lys
263 65      70      75      80
265 gaa aag tgt caa gct caa gac gtc ctc ctt tta ctg ttt gta aaa act 288
266 Glu Lys Cys Gln Ala Gln Asp Val Leu Leu Leu Leu Phe Val Lys Thr
267      85      90      95
269 gct cct gaa aac tat gat cga cgt tcc gga att aga agg acg tgg ggc 336
270 Ala Pro Glu Asn Tyr Asp Arg Arg Ser Gly Ile Arg Arg Thr Trp Gly
271      100     105     110
273 aat gaa aat tat gtt cgg tct cag ctg aat gcc aac atc aaa act ctg 384
274 Asn Glu Asn Tyr Val Arg Ser Gln Leu Asn Ala Asn Ile Lys Thr Leu
275      115     120     125
277 ttt gcc tta gga act cct aat cca ctg gag gga gaa gaa cta caa aga 432
278 Phe Ala Leu Gly Thr Pro Asn Pro Leu Glu Gly Glu Glu Leu Gln Arg
279      130     135     140
281 aaa ctg gct tgg gaa gat caa agg tac aat gat ata att cag caa gac 480
282 Lys Leu Ala Trp Glu Asp Gln Arg Tyr Asn Asp Ile Ile Gln Gln Asp
283 145     150     155     160
285 ttt gtt gat tct ttc tac aat ctt act ctg aaa tta ctt atg cag ttc 528
286 Phe Val Asp Ser Phe Tyr Asn Leu Thr Leu Lys Leu Leu Met Gln Phe
287      165     170     175
289 agt tgg gca aat acc tat tgt cca cat gcc aaa ttt ctt atg act gct 576
290 Ser Trp Ala Asn Thr Tyr Cys Pro His Ala Lys Phe Leu Met Thr Ala
291      180     185     190
293 gat gat gac ata ttt att cac atg cca aat ctg att gag tac ctt caa 624
294 Asp Asp Asp Ile Phe Ile His Met Pro Asn Leu Ile Glu Tyr Leu Gln
295      195     200     205

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/945,254

DATE: 09/21/2001

TIME: 20:48:56

Input Set : A:\Seqlist.txt

Output Set: N:\CRF3\09212001\I945254.raw

L:13 M:270 C: Current Application Number differs, Replaced Current Application No

L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date